A prototype: Reading of documents Web for visually impaired persons

Santiago S.[1], Martínez A.[2], Leija L.[3], Hernández P.[3]
[1] Dep. I.C.E.-ESIME Zacatenco, Unidad A. L. M., IPN, México
[2] Sección Computación Dep. Eléctrica, CINVESTAV-IPN, México
[3] Sección Bioelectrónica Dep. Eléctrica, CINVESTAV-IPN, México

*Abstrac*t- **This article describes a system for knowledge-based information extraction from Web-documents, reading HTML and XML structured documents using logic inference and text categorization. The system uses a multimedia speech synthesizer, into windows environment. The users could be normal persons searching filtered information without paying his whole attention, display-disable readers or visually impaired persons. In to first stage structured digital texts like public domain books (e-books) it plows read and translated to voice, their headed, name of chapters, subdivisions, words key, and paragraphs of the text are read. The user personalizes his filters using keywords and the user can move inside the text through the paragraphs and listening his contents. This project mainly increases the availability of electronic-texts for persons with a reading handicap.**

***Keywords* – Text to speech, natural language interface, information extraction**

I. INTRODUCTION

A wide necessity exists to have a system that emits synthetic voice in the Spanish language, there is a great quantity of visually disabled users that they require to be informed, there are also applications in instrumentation and in general where there is an interaction between visually disabled people and a machine or an instrument. Some technologies for disable people are available [1], like improved computer displays, screen readers, Braille displays, voice input systems and browser developments, but in the case of screen readers they are not specifically designed for Web use [2].

A voice synthesizer is used in combination with technical of artificial intelligence to develop a system of extraction of information of documents Web, to help to a better communication of the Hispanic users using their natural language. The development provides a system that emits voices in Spanish of the words contained in the paragraphs of a text in format ASCII, HTML, or XML, of internet Web pages, carrying out a concatenation of the phonemes according to the sequence of the letters of the text.

The development gives a system with emission of voices in Spanish of the words contained in the paragraph of texts in format ASCII, HTML or XML of Internet Web pages. The system defines a group of rules for the pronunciation of the words, through the selection of the phonemes that these correspond to the successions of letters of the words in Spanish. The tone of the emitted voice is according to the punctuation of the paragraph, this way it changes if it is a question, a statement or an admiration phrase.

The system has an interface based on voice, asking to the user if he requires the reading of the whole text or he prefers only the heads of the text.

II. METHODOLOGY

The system works using methods of the artificial intelligence, where the inputs are the texts and the outputs are the words spoken in Spanish. In the figure 1 the block called inference motor has in his input the structured document and executes an inference program that applies the rules stored in the block Databases of rules and then it emits a sequence of characters that they are applied to the voice synthesizer.

The system uses a general representation of the knowledge of the artificial intelligence [3], and it is represented in the following outline model:

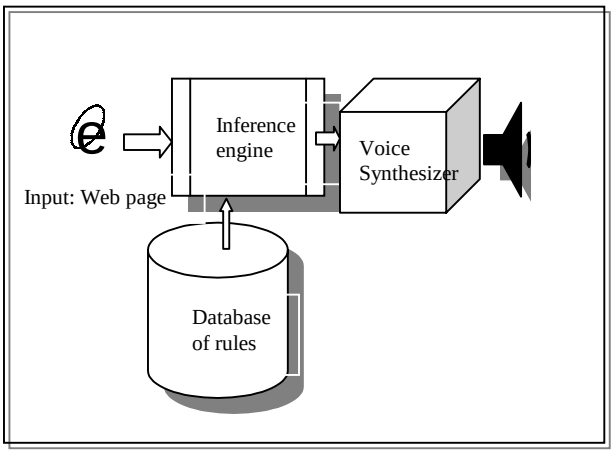

Fig 1. Outline work model.

The inference rule has the form:

Antecedent → Consequent (1)

<Word> (Sequence of letters) → phoneme(s) (2)

phoneme(s) → Spanish-Word (3)

For each word, we have:

<Word> :=<f1> <f2> <f3>... <fn> (4)

Where: f1, f2, fn, are the connected phonemes, that form the word.

The Database of rules is defined before the system is used, and it is based on the selection of the phonemes or in its individual sounds. Each word is revised to determine if the sequence of characters coincides with a sequence of an inference rule and it chains the phonemes with the previous phonemes, forming the final word.

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25OCT2001	N/A	-

Title and Subtitle A prototype: Reading of documents Web for visually impaired persons	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Dep. I.C.E.-ESIME Zacatenco, Unidad A. L. M., IPN, México	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.

Abstract

Subject Terms

Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU

Number of Pages
2

We use the synthesizer RC sys [4], it uses the technique of code lineal predictive (LPC). It has some phonemes in Spanish for the vowels and some letters, and these are:
a, e, ei, i, o, u, ñ, n, y.
A programming tool was used, denominated Developer's tools kit[4], and it is used for the control of:
Tone.
Speed.
Volume.
Punctuation.
Way text.
One program make the definition of the base of rules and another program reads the text from a file or it takes a sentence from the keyboard. It is possible to change the language without changing the code of the programs because it is only necessary to change the file that defines to the database of rules for this language. An algorithm of search of sequence of characters corresponding to sounds in the base of rules is used [5].
The user can advance of paragraph in paragraph and to return and to repeat the same paragraph. The programs were developed in language Visual C++.
The set of phonemes used in Mexican Spanish is practically a subset of phonemes used in English, this way the table 1 shows consonants and their sounds in phonemes in Spanish.

Table 1. Letters consonants and their issued phonemes.

Consonants	Sound	Consonants	Sound
b	b	ñ	ñ
c	k	p	p
d	dh	q	k
f	f	r	r
g	g	s	s
h	-	t	t
j	h	v	v
k	k	w	w
l	l	y	y
ll	y	z	s
m	m		
n	n		

III. Results

The system transforms texts with formats: ASCII, HTML or XML of Internet Web pages to voice in Spanish, with angle-American accent, because the circuit voice-synthesizer is manufactured essentially for the English language, however their voices are recognized in acceptable form and the recognition is increased when the hearing gets used to their sounds. The system has problems to accentuate the letter “a” at the end of the words, like in the words in Spanish “mamá” and “papá”. In this first stage structured digital texts like public domain books (e-books) are read and translated to voice.

The system will be being improved to complete it with the automatic searching specific information into the Web.

IV. Conclusion

The system can be used to provide an interface to give communication language natural between man and machine, and in applications prosecution information to obtain intelligent instrumentation [6]. In general the system can read structured documents used in Web-Internet, the user can select through an interface via voice the content that he wants to listen. The system is very useful to process structured documents of the Internet mainly the books type e-book, where the visually weak or visually impaired user moves through them listening the paragraphs. The system can also be used for reading of texts in English with an excellent accent. The system can be used to provide an interface of communication in language natural man-machine and to obtain intelligent instrumentation. The system can also be used for treatment of texts in English with an excellent accent.

References

[1] J. J. Lázaro, “Helping the Web help the disabled,” *IEEE Spectrum*, vol. 36 pp. 54-59, March 1999.
[2] M. Zajicek, and C. Powell, “Enabling visually impaired people to use the Internet,” Computers in the Service of Mankind,” *IEE Colloquium Digest* No.1997/117, pp. 11/1-11/3, March 1997.
[3] N. J. Nilsson, *Principles of Artificial Intelligence*, Tioga Publishing, 1980.
[4] RC Systems Inc, “Double Talk PC/LT user’s Manual,” 1996.
[5] A. M. Martínez-Enríquez, and Sánchez-Salguero H., “The artificial intelligence and some of their applications,” volume II, voice prosecution, Technician report CINVESTAV_IPN, México, 1987.
[6] A. Ramírez-Treviño, “Sistema digital de reconocimiento de palabras aisladas para generar comandos de activación a aparatos electrónicos,” Master’ thesis, CINVESTAV-IPN, México, 1997.